THÈME VII

Introduction au raisonnement en présence de connaissances **incomplètes et/ou incohérentes**

Victor David et MC Lagasquie-Schiex

UNIVERSITÉ TOULOUSE III PAUL SABATIER
Université de Toulouse

Plan

I. Introduction

II. Raisonnement sous l’Hypothèse de Monde Clos

III. Raisonnement plausible et traitement de l’incohérence

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

Plan

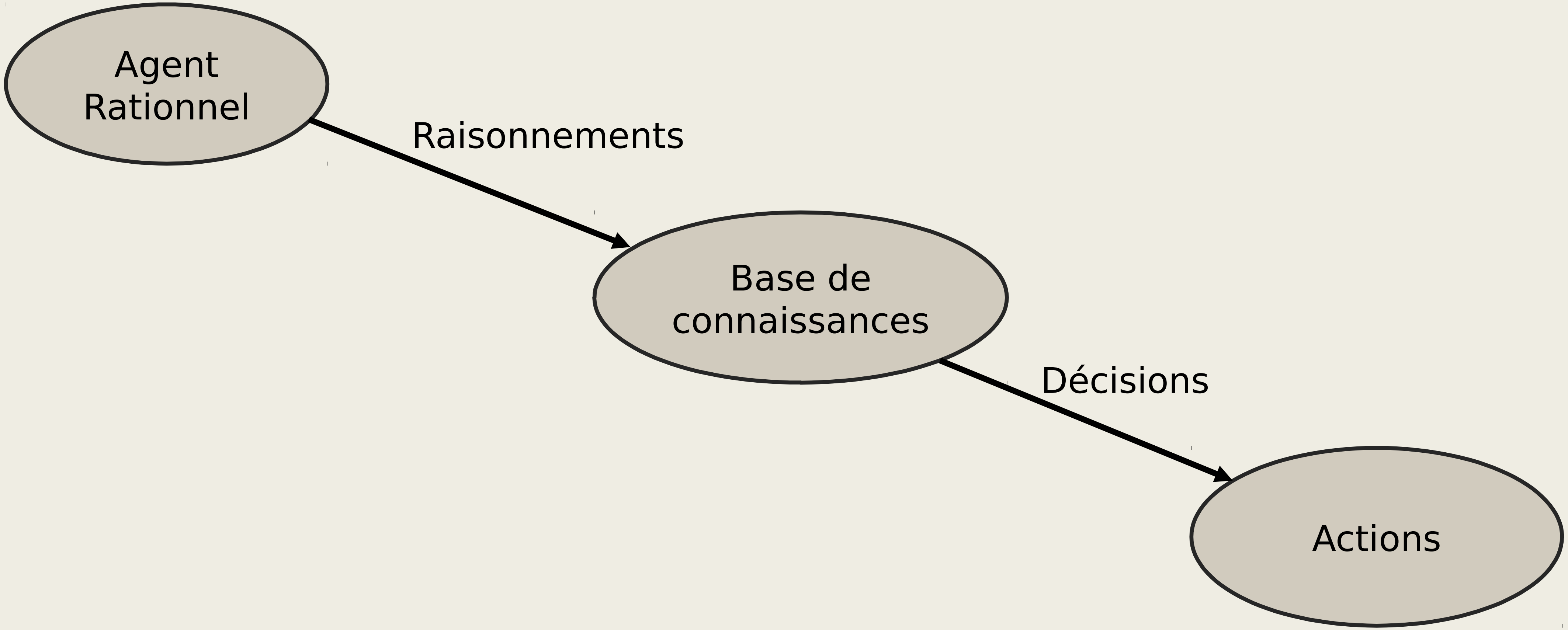
I. Introduction
   a) *Raisonnement et connaissances de sens commun*
   b) *Raisonnement révisable*

II. Raisonnement sous l'Hypothèse de Monde Clos

III. Raisonnement plausible et traitement de l'incohérence

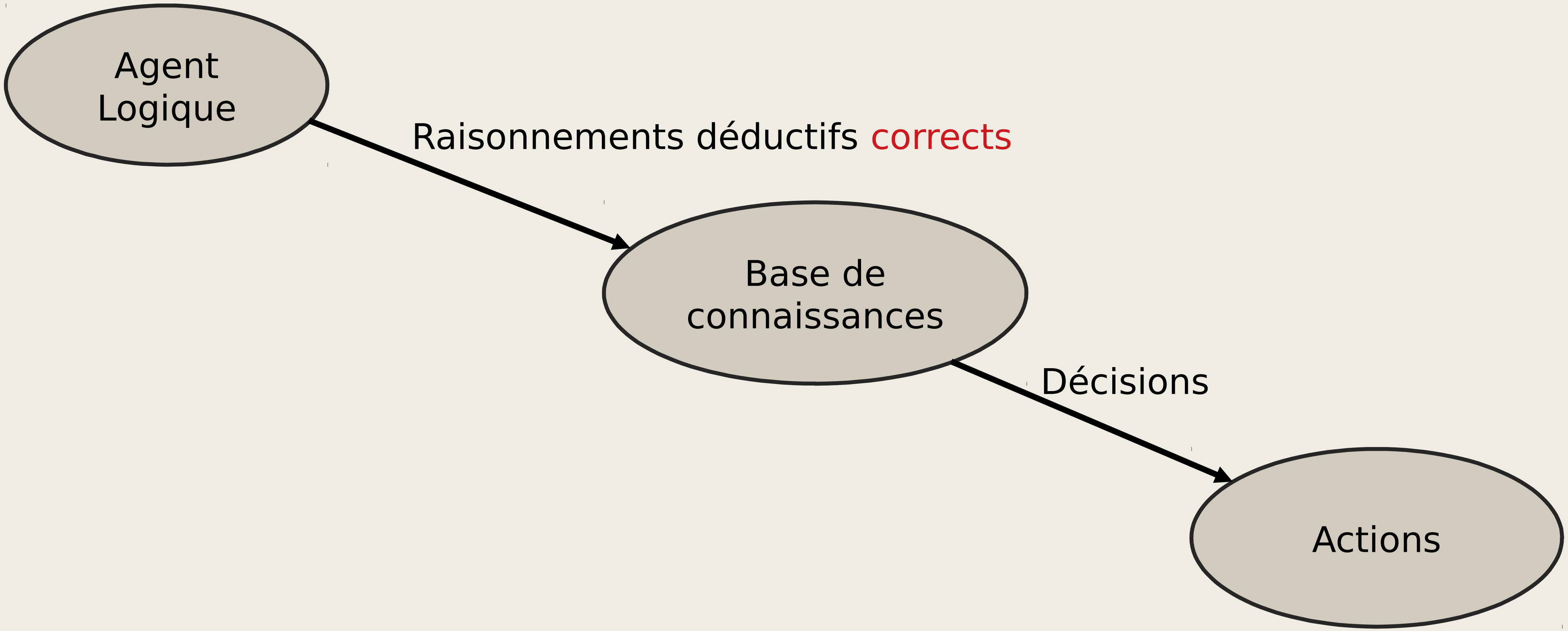
I. Introduction

a) Raisonnement et connaissances de sens commun



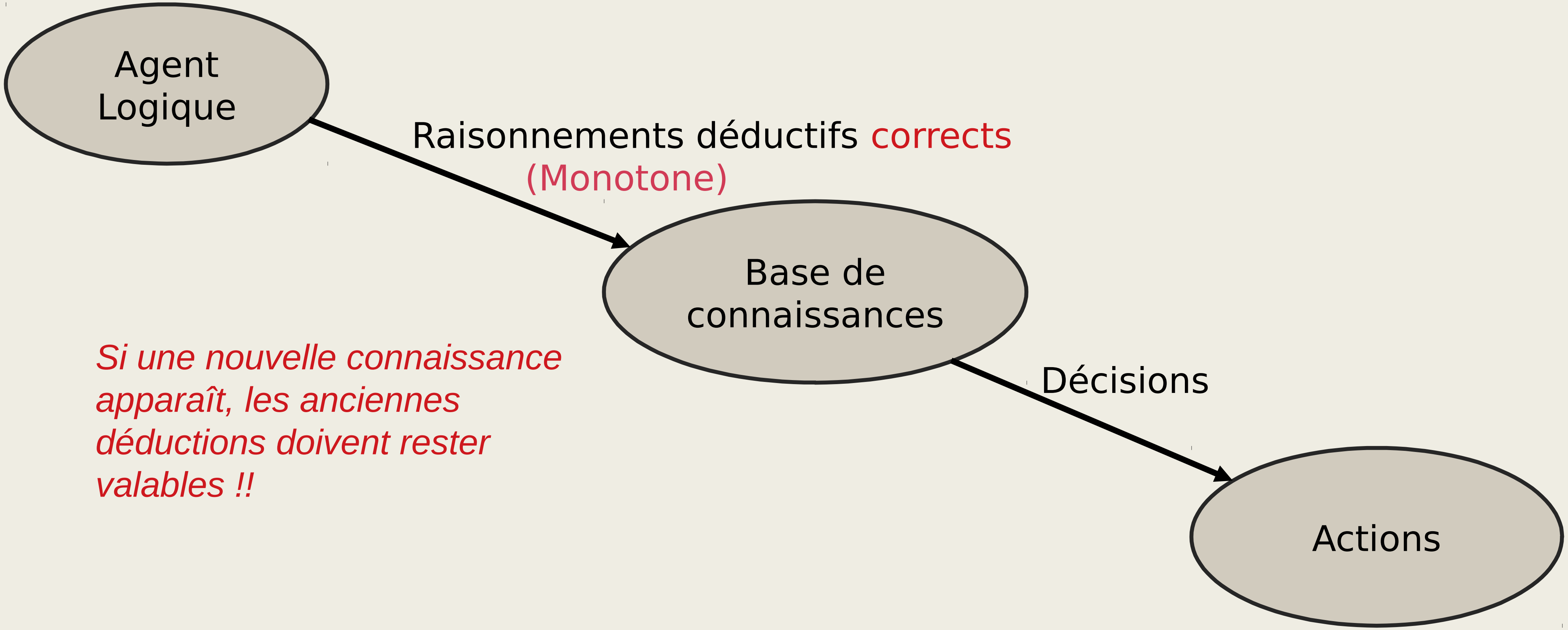
I. Introduction

a) Raisonnement et connaissances de sens commun



I. Introduction

a) Raisonnement et connaissances de sens commun



Si une nouvelle connaissance apparaît, les anciennes déductions doivent rester valables !!

I. Introduction

a) Raisonnement et connaissances de sens commun

Objectif, définir des agents capables de traiter des **connaissances de sens commun** :

- connaissances incomplètes
- partiellement incohérentes

Conclusions simplement plausibles et sujettes à révisions → Raisonnement **révisable**

I. Introduction

b) Raisonnement révisable

Deux types de problèmes différents à prendre en compte :

- Il nous manque des informations (**incomplétude**)
- On a trop d'informations (**inconsistance**)

A chaque type de problème, une technique !

Mais toujours la **non-monotonie :**

Un raisonnement révisable n'est pas correct :

$$A \models C \quad \text{mais} \quad A \cup B \not\models C$$

I. Introduction

b) Raisonnement révisable

En présence d’informations **incomplètes**, on va faire des **hypothèses** sur l’état des facteurs inconnus

Les **hypothèses** seront révisables par l’apparition de nouveaux faits

Exemple : (on ne sait rien sur la normalité de Titi)

Un oiseau normal vole et Titi est un oiseau $\models$ Titi vole (on a supposé que Titi était normal)

Titi est une autruche et les autruches ne sont pas normales $\models$ Titi ne vole pas

(on a révisé notre croyance sur la normalité de Titi)

I. Introduction

b) Raisonnement révisable

En présence d'**inconsistance**, on considère que nos connaissances reflètent plusieurs images du monde incohérentes deux à deux

Un raisonnement non monotone utilise une relation d'inférence où les conclusions ne sont pas vérifiées dans tous les modèles des prémisses :
relation de **conséquence affaiblie**

→ **Pluri-extensionalité** : plusieurs ensembles de conclusions plausibles

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

I. Introduction

b) Raisonnement révisable

Exemple :

BC = { Un oiseau vole, Titi est un oiseau,
Titi est une autruche, les autruches ne volent pas }

Plusieurs images du monde sont possibles :

- Celle où Titi vole car c'est un oiseau
- Celle où Titi ne vole pas car c'est une autruche

Par contre, on veut quand même éviter que des conclusions mutuellement inconsistantes soient inférées conjointement !
(dans une même image du monde)

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

Plan

I. Introduction

II. Raisonnement sous l’Hypothèse de Monde Clos (complétude)
   a) *Définitions*
   b) *Propriétés et limitations*
   c) *Hypothèse restreinte de monde clos*
   d) *Conventions supplémentaires*

III. Raisonnement plausible et traitement de l’incohérence

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

II. Raisonnement sous l’Hypothèse de Monde Clos

a) Définitions

Idée : le raisonnement à partir de connaissances incomplètes peut exploiter un principe de **minimisation de l’information implicite** pour compléter la base de connaissances

Définition : l’Hypothèse de Monde Clos de Reiter (HMC) énonce que toute information élémentaire positive qui **ne peut être déduite** (de façon classique) **est fausse**

« Q étant une formule atomique sans variable du langage des prédicats, si Q n’est pas déductible de la base de connaissances, alors on peut inférer la négation de Q »

II. Raisonnement sous l'Hypothèse de Monde Clos

a) Définitions

Définition : soit BC un ensemble de formules d'un langage L.
L'ensemble de toutes les **conséquences logiques** de BC est la **théorie logique** associée à BC notée **Th(BC)**

Propriété : une théorie logique T est **complète** ssi pour toute formule atomique **de base** (c-à-d sans variable) du langage, elle ou sa négation est présente dans la théorie

Exemple : $BC = \{ P(a), P(b), P(a) \rightarrow Q(a) \}$

Th(BC) n'est pas complète car elle ne contient ni $Q(b)$ ni $\neg Q(b)$

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

II. Raisonnement sous l'Hypothèse de Monde Clos

a) Définitions

L'hypothèse de monde clos permet de compléter une théorie non complète

Définition : BCs est l'ensemble des littéraux négatifs de base qui correspondent aux littéraux positifs de base n'appartenant pas à Th(BC), et qui n'appartiennent pas déjà à Th(BC)

Définition : $HMC_L(BC) = Th(BC \cup Bcs)$ avec L le langage sur lequel est construit BC

Exemple : $BC = \{ P(a), P(b), P(a) \rightarrow Q(a) \}$ avec $L = \{a,b\}$

Alors $BCs = \{ \neg Q(b) \}$ et

$HMC_L(BC) = \{P(a), P(b), Q(a), \neg Q(b) \}$

Université Toulouse III Paul Sabatier — Université de Toulouse

II. Raisonnement sous l'Hypothèse de Monde Clos

b) Propriétés et limitations

Propriété [Préservation de la consistance] : Soit BC un ensemble consistant de formules. La théorie $HMC_L(BC)$ est **inconsistante** ssi il existe des formules atomiques de base positives $\Phi_1, \dots, \Phi_n$ telles que :

- $\forall i \in \{ 1, \dots, n \}$, $BC \nvdash \Phi_i$
- $BC \vdash \Phi_1 \vee \dots \vee \Phi_n$

Exemple : BC = { P(a) v P(b) } avec L = {a, b}

Alors $BCs = \{ \neg P(a), \neg P(b) \}$ et

$HMC_L(BC) = \{ \neg P(a), \neg P(b), P(a) \vee P(b) \}$ est inconsistante

Corollaire : si BC est un ensemble consistant de clauses de Horn, $HCM_L(BC)$ est consistant

II. Raisonnement sous l'Hypothèse de Monde Clos

b) Propriétés et limitations

La définition du langage logique utilisée est **très importante** :

Exemple : BC = { P(a), Q(b), P(x) v Q(x) }

Si L = { a, b }
Alors HMC_L(BC) est consistante

langage implicite : HMC(BC)

Si L = { a, b, c }
Alors HMC_L(BC) est inconsistante

langage explicite : HMC_L(BC)

II. Raisonnement sous l’Hypothèse de Monde Clos

b) Propriétés et limitations

Limitation de l’implémentation du « non déductible de »

Pour prouver le fait que BC n’infère pas une formule, nous nous heurtons à un problème :

la logique des prédicats est seulement **semi-décidable**

(si la preuve de P existe, on la trouvera,
mais si elle n’existe pas on peut ne jamais s’arrêter de chercher)

Exemple de la « négation par l’échec » en programmation logique qui peut échouer car la preuve pour P est infinie

II. Raisonnement sous l'Hypothèse de Monde Clos

b) Propriétés et limitations

Limitation de l'implémentation du « non déductible de »

La dérivation de ¬P selon le principe de Négation par échec consiste en une recherche *exhaustive* et *sans succès* d'une preuve de P

HMC	Preuve de P	Négation par Echec
$BC \vdash P$: $\neg P \notin BCs$	Arrêt avec Succès	On ne rajoute pas ¬P
$BC \nvdash P$: $\neg P \in BCs$	Arrêt avec Echec	On rajoute ¬P
	???	**On ne rajoute pas ¬P**

Exemple : BC = { $(P(x, y) \wedge Q(y, z)) \rightarrow P(x, z)$ }
P(a, b) n'est pas classiquement prouvable à partir de BC.

Cependant, si l'on cherche à prouver P(a, b) en Prolog :
l'arbre de recherche est infini !!

II. Raisonnement sous l’Hypothèse de Monde Clos

c) Hypothèse restreinte de monde clos

Comment répondre au risque d’obtenir une HMC inconsistante ?

On peut affaiblir l’hypothèse de monde clos en ne l’appliquant qu’à un sous ensemble de prédicats

Exemple : BC = { $P(x) \rightarrow Q(x)$,
$R(b) \vee Q(b)$,
$P(a)$ }

Si l’on restreint l’hypothèse de monde clos à Q :

HMC(BC, {Q}) = { $P(a)$, $Q(a)$, $R(b) \vee Q(b)$, $\neg Q(b)$, $R(b)$ }

Rq : Sans restriction, HMC(BC) est inconsistante (à cause de R)

II. Raisonnement sous l'Hypothèse de Monde Clos

d) Conventions supplémentaires

(en général associées à l'utilisation de l'HMC)

Hypothèse de Fermeture du Domaine (Domain Closure Assumption) (**DCA**) : limite le domaine des variables aux seuls objets qui peuvent être dénotés par des termes de base du langage

Par exemple, si on n'a pas de fonction et que l'ensemble des constantes = $\{c_1, \dots, c_n\}$
Alors $\forall x\ ((x = c_1) \lor \dots \lor (x = c_n))$

Hypothèse d'Unicité des Noms (Unique Name Assumption) (**UNA**) : les termes de base dont on ne peut pas prouver qu'ils sont égaux, sont supposés différents

Par exemple, si on n'a pas de fonction et que l'ensemble des constantes = $\{c_1, \dots, c_n\}$
Alors $\forall i, j \in \{1, \dots, n\}$ et $i \neq j$, $\neg(c_i = c_j)$ (sauf si $(c_i = c_j)$ est prouvable)

Exercices

TD 7 : exos 1, 2 et 3

Plan

I. Introduction

II. Raisonnement sous l'Hypothèse de Monde Clos

III. Raisonnement plausible et traitement de l'incohérence
   a) *Introduction*
   b) *Restauration de cohérence*
   c) *Conséquence forte et inférence dans les modèles préférés*

III. Raisonnement plausible et traitement de l’incohérence

a) Introduction

Il existe des situations variées dans lesquelles les connaissances disponibles sont incohérentes :

- Plusieurs experts ou capteurs fournissent des données conflictuelles
- Plusieurs règles sont applicables au contexte courant mais ont des conclusions contradictoires
- Les objectifs de plusieurs agents sont incompatibles

En présence d’inconsistance, la déduction logique classique est inutilisable (elle permet de déduire toute formule du langage). Lorsqu’on n’a pas la possibilité d’acquérir de nouvelles informations pour corriger la base, on doit raisonner de manière à réduire l’ensemble des conclusions

approche consistant à affaiblir la relation de conséquence.

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

L'idée est de sélectionner des sous-bases cohérentes (dans un sens à définir) de la base initiale de manière à pouvoir utiliser des méthodes d'inférence classique sur ces sous-bases.

Base syntaxique :

- Nous considérons des bases de connaissances classiquement inconsistantes (i.e. qui n'ont pas de modèle consistant).
- Nous appellerons croyance un élément de connaissance qu'on accepte de rejeter lors du processus de raisonnement.
- Chaque croyance est une entité indivisible, que l'on acceptera ou rejettera telle quelle lors du processus de raisonnement.
- La base de croyances n'est pas supposée consistante, ni fermée pour la conséquence sémantique (c-a-d. il peut y avoir des variables). En conséquence, deux bases logiquement équivalentes ne seront pas nécessairement traitées de la même façon.

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

Exemple : la base $X = \{A, B\}$ ne sera pas traitée de la même manière que
la base $Y = \{A \wedge B\}$

Rq : $X \equiv Y$ (logiquement équivalent) car $Th(X) = Th(Y)$

Soit l'interprétation 1, $I_1[A]$ = Vrai et $I_1[B]$ = Vrai	
$M_1[X] = \{ I_1[A], I_1[B] \}$ est un modèle de X et on peut inférer A et B	$M_1[Y] = \{ I_1[A], I_1[B] \}$ est un modèle de Y et on peut inférer A et B

Soit l'interprétation 2, $I_2[A]$ = Vrai et $I_2[B]$ = Faux	
$M_2[X] = \{ I_2[A], I_2[B] \}$ n'est pas un modèle de X car $I_2[B]$ = Faux, mais on peut inférer A	$M_2[Y] = \{ I_2[A], I_2[B] \}$ n'est pas un modèle de Y et on ne peut rien inférer

A et B sont syntaxiquement dépendants dans Y

III. Raisonnement plausible et traitement de l’incohérence

b) Restauration de cohérence

Schéma de base : BC = (N, H)

Considérons une base de connaissances BC composée de deux parties :

- N (le noyau) est un ensemble de connaissances sûres, supposé consistant. Ces connaissances ne devront pas être remises en question lors du raisonnement.
- H est un ensemble de croyances (i.e. informations sujettes à révision lors du raisonnement). L’ensemble H peut être logiquement inconsistant avec le noyau.

Cette inconsistance sera traitée en exploitant les sous-ensembles N-consistants de la base de connaissances.

Dans ce qui suit, nous manipulerons des croyances de base, obtenues par instanciation des croyances sur le domaine de Herbrand (ensemble des termes de base) défini par le noyau.

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

Nous appellerons scénario un sous-ensemble de croyances consistant avec N et maximal pour l'inclusion ensembliste (on dit aussi sous-ensemble maximal N-consistant).

Définition: $S \subseteq H$ est un scénario de BC ssi :

(1) L'union de S et N est consistant (i.e. $N \cup S$ est consistant)

(2) S est maximal (pour l'inclusion ensembliste) parmi les sous-ensembles de H vérifiant (1).

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

On s'intéresse à l'ensemble des conséquences logiques de l'union d'un scénario et de l'ensemble N. Cet ensemble est appelé extension de BC dans la suite.

Définition: Soit S un scénario de BC. $E = Th(N \cup S)$ est une extension de BC.

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

Il existe généralement plusieurs scénarios, donc plusieurs extensions.
Cette multiplicité conduit à différentes relations de conséquence.

Définition:

- Q est une conséquence forte de BC ssi Q appartient à *toute* extension de BC.
- Q est une conséquence faible de BC ssi Q appartient à *au moins une* extension de BC.
- Q est une conséquence argumentative de BC ssi Q est une conséquence faible de BC mais la négation de Q n'est pas conséquence faible de BC.

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

Exemple : Considérons la base de connaissances suivante BC = (N, H) avec :

N = { $O(Titi), A(Zoé), A(x) \rightarrow O(x)$ }

H contient deux croyances libres (formules ouvertes) :

H = { $A(x) \rightarrow \neg V(x), O(x) \rightarrow V(x)$ }

Chaque croyance libre correspond à deux croyances de base (1 avec Titi et 1 avec Zoé). On obtient deux scénarios :

S1 = { $A(Titi) \rightarrow \neg V(Titi), A(Zoé) \rightarrow \neg V(Zoé), O(Titi) \rightarrow V(Titi)$ }

S2 = { $A(Titi) \rightarrow \neg V(Titi), O(Zoé) \rightarrow V(Zoé), O(Titi) \rightarrow V(Titi)$ }

Rq : O(Titi) étant toujours vrai, on obtient V(Titi) à vrai (et donc A(Titi) à faux).
Et pour Zoé, on a A(Zoé) à vrai et donc O(Zoé) à vrai. Par contre, suivant le scénario, on aura soit V(Zoé) à vrai, soit V(Zoé) à faux

La conclusion V(Titi) est une conséquence forte de BC, alors que ¬V(Zoé) est seulement conséquence faible.

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

III. Raisonnement plausible et traitement de l'incohérence

b) Restauration de cohérence

Remarque : Le critère de maximalité pour l'inclusion n'est généralement pas assez sélectif. Un critère permettant une meilleure discrimination est la maximalité pour la cardinalité.

Exemple : BC' = (N', H') avec :

N' = { O(titi), A(Zoé), A(x) → O(x), O(x) → AdesAiles(x) }
H' = { A(x) → ¬V(x), O(x) → V(x), AdesAiles(x) → V(x) }

Avec la maximalité pour l'inclusion on obtient toujours deux scénarios :

S1' = { A(Titi) → ¬V(Titi), A(Zoé) → ¬V(Zoé), O(Titi) → V(Titi), AdesAiles(Titi) → V(Titi) }
S2' = { A(Titi) → ¬V(Titi), O(Zoé) → V(Zoé), O(Titi) → V(Titi), AdesAiles(Titi) → V(Titi), AdesAiles(Zoé) → V(Zoé) }

Avec la maximalité pour la cardinalité, on obtient 1 seul scénario :

S2' = { A(Titi) → ¬V(Titi), O(Zoé) → V(Zoé), O(Titi) → V(Titi), AdesAiles(Titi) → V(Titi), AdesAiles(Zoé) → V(Zoé) }

UNIVERSITÉ TOULOUSE III PAUL SABATIER Université de Toulouse

III. Raisonnement plausible et traitement de l'incohérence

c) Conséquence forte et inférence dans les modèles préférés

Soit BC = (N, H). Rappelons que l'ensemble N est supposé consistant. Il admet donc au moins un modèle. L'idée est d'utiliser H pour comparer les modèles de N.

Définition : Soit M et M' deux interprétations de (N ∪ H). M ≤ M' ssi l'ensemble des croyances (de base) vraies dans M est un sous-ensemble de l'ensemble des croyances (de base) vraies dans M'.

Remarque : cette relation est réflexive et transitive.

UNIVERSITÉ TOULOUSE III PAUL SABATIER

III. Raisonnement plausible et traitement de l'incohérence

c) Conséquence forte et inférence dans les modèles préférés

Définition : Une interprétation M est un modèle préféré de N ssi M est un modèle de N et il n'existe pas de M' modèle de N tel que M ≤ M' et pas M' ≤ M (c'est-à-dire M < M').

Propriété : Soit M un modèle de N. M est un modèle préféré de N ssi l'ensemble des croyances de base vraies dans M est un scénario de BC.

Corollaire : Q est une conséquence forte de BC ssi Q est vraie dans tous les modèles préférés de N.

III. Raisonnement plausible et traitement de l'incohérence

c) Conséquence forte et inférence dans les modèles préférés

Exemple : BC = (N, H) avec
N = { O(Titi), A(Zoé), A(x) → O(x) } et
H = { A(x) → ¬V(x), O(x) → V(x) }

Il existe 8 modèles de N parmi lesquels on trouve 2 modèles préférés :

	O(Titi)	O(Zoé)	A(Titi)	A(Zoé)	V(Titi)	V(Zoé)	
M1	Vrai	Vrai	Faux	Vrai	Vrai	Vrai	S1
M2	Vrai	Vrai	Faux	Vrai	Vrai	Faux	S2

Les conséquences fortes

Exercices

TD 7 : exos 4 et 5